

Foundation of Cryptography, Lecture 10

Pseudorandom Generator from One-Way Functions

Handout Mode

Iftach Haitner, Tel Aviv University

Tel Aviv University.

May 27, 2014

Section 1

Entropy

Different measures of entropy

Let X be a random variable over \mathcal{U} and let $X(x) = \Pr_X[x]$.

- **Support**: $\text{Supp}(X) := \{x \in \mathcal{U} : X(x) > 0\}$.
- **Max entropy**: $H_0(X) = \log |\text{Supp}(X)|$.
- **Sample entropy**: For $x \in \text{Supp}(X)$: $H_X(x) = \log \frac{1}{X(x)}$.¹
- **Shannon entropy**: $H(X) = \sum_{x \in \mathcal{U}} X(x) \cdot H_X(x) = \mathbb{E}_X [H_X(x)]$
- **Collision probability**: $\text{CP}(X) = \sum_{x \in \mathcal{U}} X(x)^2 = \Pr_{x, x' \leftarrow X} [x = x']$
- **Renyi entropy**: $H_2(X) = \log \frac{1}{\text{CP}(X)}$
- **Min entropy**: $H_\infty(X) = \min_{x \in \text{Supp}(x)} \{H_X(x)\}$

It holds that $0 \leq H_\infty(X) \leq H_2(X) \leq H(X) \leq H_0(X)$.

Equality iff X is **uniform**.

¹All logarithmic are on base 2.

Conditional Entropy

Given two random variable X and Y , the conditional Shannon entropy of X given Y is defined as

$$H(X | Y) = \mathbb{E}_{y \leftarrow Y} [H(X | Y = y)]$$

Example: let $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ be a 2^k regular function. Let X be uniform over $\{0, 1\}^n$ and let $Y = f(X)$. Then

$$H(X | Y) = \mathbb{E}_{y \leftarrow Y} [\log 2^k] = k.$$

Flattening Shannon entropy

Lemma 1

Let X be a rv over \mathcal{U} , let $t \in \mathbb{N}$ and let $\varepsilon > 0$. Then \exists rv Z that is $(\varepsilon + 2^{-t})$ -close to X^t , and $H_\infty(Z) \geq t \cdot H(X) - O(\sqrt{t \cdot \log(1/\varepsilon)} \cdot \log(|\mathcal{U}| \cdot t))$.

Proof: ?

Pairwise independent hashing

Definition 2 (pairwise independent function family)

A function family $\mathcal{H} = \{h: \{0, 1\}^n \mapsto \{0, 1\}^m\}$ is **pairwise independent**, if $\forall x \neq x' \in \{0, 1\}^n$ and $y, y' \in \{0, 1\}^m$, it holds that $\Pr_{h \leftarrow \mathcal{H}}[h(x) = y \wedge h(x') = y'] = 2^{-2m}$.

Example $\mathcal{H} = \{(A, b) \in \{0, 1\}^{m \times n} \times \{0, 1\}^m\}$ with $(A, b)(x) = A \times x + b$.

We identify functions with their description, and assume wlg. that (the description of) a random element from \mathcal{H} is a uniform string.

Lemma 3 (leftover hash lemma)

Let X be a rv over $\{0, 1\}^n$ with $H_2(X) \geq k$ and let $\mathcal{H} = \{h: \{0, 1\}^n \mapsto \{0, 1\}^m\}$ be pairwise independent, then

$$SD((H, H(X)), (H, U_m)) \leq 2^{(m-k-2)/2},$$

where H is uniformly distributed over \mathcal{H} and U_m is uniformly distributed over $\{0, 1\}^m$.

Computational notions of entropy

Definition 4

A random variable has **pseudoentropy** at least k , if it is computationally indistinguishable from a RV Y with $H(Y) \geq k$.

Pseudo min/Reiny -entropy are analogously defined.

- Examples
- Repeated sampling

Section 2

PRG from Regular OWF

PRG from Regular OWF

Definition 5

Given a function $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ and function family

$\mathcal{H}: \{0, 1\}^n \mapsto \{0, 1\}^m$, let $g = g(f, \mathcal{H}): \mathcal{H} \times \{0, 1\}^n \mapsto \mathcal{H} \times \{0, 1\}^n \times \{0, 1\}^m$ be defined by $g(h, x) = (g(x), h, h(x))$.

In case f and \mathcal{H} are function families, we let $g(f, \mathcal{H}) = \{g(f_n, \mathcal{H}_n)\}_{n \in \mathbb{N}}$.

Claim 6

Let f be a $2^{k=k(n)}$ -regular OWF, $\mathcal{H} = \{\mathcal{H}_n: \{0, 1\}^n \mapsto \{0, 1\}^{m(n)=k(n)+\log n}\}$ be efficient family of pairwise independent hash function family, and let $g = g(f, \mathcal{H})$. Then

- 1 $H(g(U_n, H_n)) \geq n + H(H_n) - \frac{1}{n}$, where H_n is uniform over \mathcal{H}_n .
- 2 g is one-way.

g has high entropy

$$\begin{aligned}\text{CP}(g(U_n, H_n)) &:= \Pr_{w, w' \leftarrow \{0,1\}^n \times \mathcal{H}_n} [g(w) = g(w')] \\ &= \Pr_{h, h' \leftarrow \mathcal{H}_n} [h = h'] \cdot \Pr_{x, x' \leftarrow \{0,1\}^n} [f(x) = f(x')] \\ &\quad \cdot \Pr_{h \leftarrow \mathcal{H}_n; x, x' \leftarrow \mathcal{Z}^n} [h(x) = h(x') \mid f(x) = f(x')] \\ &= \text{CP}(H_n) \cdot \text{CP}(f(U_n)) \cdot (2^{-k} + (1 - 2^{-k}) \cdot 2^{-m}) \\ &\leq \text{CP}(H_n) \cdot \text{CP}(f(U_n)) \cdot (2^{-k} + 2^{-m}) \\ &\leq \text{CP}(H_n)(2^{-n} + 2^{-n - \log n}) = \text{CP}(H_n) \cdot \text{CP}(U_n) \cdot (1 + \frac{1}{n}).\end{aligned}$$

Hence, $H_2(g(U_n, H_n)) \geq H_2(\mathcal{H}_n) + H_2(U_n) + \log \frac{1}{1 + \frac{1}{n}} \geq H(H_n) + n - \frac{1}{n}$.

Thus, $H(g(U_n, H_n)) \geq H(H_n) + n - \frac{1}{n}$.

g is one-way

Assume g is not one-way and let A be a PPT that inverts g w.p $1/p(n)$, for some $p \in \text{poly}$, for infinitely many n 's.

The following algorithm inverts f with non-negligible probability.

Let $t = t(n) = k(n) - 2 \lceil \log(p(n)) \rceil$.

Algorithm 7 (B)

Input: $y \in \{0, 1\}^n$.

Sample $h \leftarrow \mathcal{H}_n$ and $z \leftarrow \{0, 1\}^t$, and return $D(y, h, z)$

Algorithm 8 (D)

Input: $y \in \{0, 1\}^n$, $h \in \mathcal{H}_n$ and $z_1 \in \{0, 1\}^t$.

For all $z_2 \in \{0, 1\}^{m-t}$:

- 1 Let $(x, h) \leftarrow A(y, h, z)$.
- 2 If $f(x) = y$, return x .

$$\Pr_{x \leftarrow \{0,1\}^n; h \leftarrow \mathcal{H}_n} [D(f(x), h, h(x)_{1,\dots,t}) \in f^{-1}(f(x))] \geq \frac{1}{p(n)} \quad (1)$$

g is one-way, cont.

By the leftover hash lemma(?)

$$\text{SD}((f(x), h, h(x)_1, \dots, t)_{x \leftarrow \{0,1\}, h \leftarrow \mathcal{H}_n}, (f(x), h, U_t)_{x \leftarrow \{0,1\}, h \leftarrow \mathcal{H}_n}) \leq \frac{1}{2p(n)} \quad (2)$$

Hence,

$$\Pr_{x \leftarrow \{0,1\}^n} [\mathbf{B}(f(x)) \in f^{-1}(f(x))] \geq \frac{1}{p(n)} - \frac{1}{2p(n)} = \frac{1}{2p(n)}.$$

The generator

Claim 9

Let $f: \{0, 1\}^n \mapsto \{0, 1\}^m$ be a OWF with $H(f(U_n)) \geq n - \frac{1}{2}$, and let b be an hardcore predicate for f . Then $g(x) = f(x) \circ b(x)$ has pseudoentropy $n + \frac{1}{2}$.

Proof: ?

We call such g a **pseudo-entropy** generator.

Claim 10

The function $g^{n^2}(x_1, \dots, x_{n^2}) = g(x_1), \dots, g(x_{n^2})$ has **pseudo min-entropy** $n(n + \frac{1}{2}) - O(\sqrt{n \log^2 n \cdot \log(n^2)}) \geq n^2 + n/2 - O(n^{2/3})$.

Proof: by the flattening lemma, taking $\varepsilon = 2^{-\log^2 n}$ and $t = n$.

Claim 11

Let $\mathcal{H}: \{0, 1\}^{n^2+n} \mapsto \{0, 1\}^{n^2+n/4}$ be an efficient pairwise hash function, then $G: \{0, 1\}^{n^2} \times \mathcal{H}_n$ defined by $G(x_1, \dots, x_{n^2}, h) = (h, h(g^{n^2}(x_1, \dots, x_{n^2})))$, is a PRG.

Proof: by the leftover hash lemma

Section 3

PRG from any OWF

Inefficient construction

Definition 12

Given a function $f: \{0, 1\}^n \mapsto \{0, 1\}^m$ and $x \in \{0, 1\}^n$, let

$$d_f(x) = \lceil \log(|f^{-1}(f(x))|) + \log n \rceil.$$

Given $\mathcal{H}: \{0, 1\}^n \mapsto \{0, 1\}^{n+\log n}$, let

$g = g(f, \mathcal{H}): \mathcal{H} \times \{0, 1\}^n \mapsto \mathcal{H} \times \{0, 1\}^n \times \{0, 1\}^{n+\log n}$ be defined by
 $g(h, x) = (f(x), h, h(x)_1, \dots, h(x)_{d_f(x)}, 1^{n+\log n - d_f(x)}).$

Claim 13

Let f be a OWF, $\mathcal{H} = \{\mathcal{H}_n: \{0, 1\}^n \mapsto \{0, 1\}^{n+\log n}\}$ be efficient family of pairwise independent hash function family, and let $g = g(f, \mathcal{H})$. Then

- 1 $H(g(U_n, H_n)) \geq n + H(H_n) - \frac{1}{n}$, where H_n is uniform over \mathcal{H}_n .
- 2 Assume d_f is poly-time computable, then g is a one-way function.

Proof:

Hence, if d_f is poly-time computable, then building a PRG from f follows the same lines we used for regular OWF.

Should we expect d_f to be poly-time computable?

Efficient construction, first approach

Definition 14

For $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ and $\mathcal{H} = \{h: \{0, 1\}^n \mapsto \{0, 1\}^{n+\log n}\}$, let $g = g(f, \mathcal{H}): \mathcal{H} \times [n] \times \{0, 1\}^n \mapsto \mathcal{H} \times [n] \times \{0, 1\}^n \times \{0, 1\}^{n+\log n}$ be defined by $g(h, i, x) = f(x), h, i, h(x)_{1, \dots, i+\log n}, 1^{n+\log n-i}$.

Claim 15

Assume f is OWF and that \mathcal{H} is the Matrix-based pairwise-independent hash functions. Then the **pseudo** Shannon-entropy of $g(H_n, I_n, U_n)$, where I_n is uniform over $[n]$, is larger by at least $1/n$ than its (real) Shannon entropy.

We call such g a **false-pseudoentropy** generator.

Proof: Define

$$g'(h, i, x) = \begin{cases} f(x), h, i, h(x)_{1, \dots, i+\log n-1}, U, 1^{n+\log n-i}, & i = d_f(x) \\ g(h, i, x), & \text{otherwise.} \end{cases}$$

Claim 16

- ① $g(H_n, I_n, U_n) \approx_c g'(H_n, I_n, U_n)$
- ② $H(g'(H_n, I_n, U_n)) - H(g(H_n, I_n, U_n)) \geq 1/n$

False-pseudoentropy generator to PRG

- 1 Using repetition convert the Shannon pseudoentropy of the output of g into min pseudoentropy.

Problem: g' is not efficiently computable, and thus $g'(H_n, I_n, U_n)$ is not efficiently samplable

- 2 “extract” this min-entropy, and also the (real) min-entropy “left” in the inputs.

Very complicated an inefficient construction. Seed length of PRG is $\Theta(n^8)$.

Efficient construction, second approach

Definition 17

For $f: \{0, 1\}^n \mapsto \{0, 1\}^n$, and the Matrix-based pairwise-independent hash functions $\mathcal{H} = \{h: \{0, 1\}^n \mapsto \{0, 1\}^{n+\log n}\}$, let $g: \mathcal{H} \times \{0, 1\}^n \mapsto \mathcal{H} \times \{0, 1\}^n \times \{0, 1\}^{n+\log n}$ be defined by $g(h, x) = (f(x), h, h(x))$.

But g is **invertible** and thus its output pseudoentropy is as large as its real entropy.(?)

Right, but not in the eyes of an **online observer**.

Next-block pseudoentropy generator

Definition 18 (next-block pseudoentropy)

$X = (X_1, \dots, X_m)$ has **next-block pseudoentropy at least k** , \exists rv $Y = (Y_1, \dots, Y_m)$, (jointly distributed with X), such that:

- 1 $\forall i, (X_1, X_2, \dots, X_{i-1}, X_i) \approx_c (X_1, X_2, \dots, X_{i-1}, Y_i)$.
- 2 $\sum_i H(Y_i | X_1, \dots, X_{i-1}) \geq k$.

Quantitative generalization of unpredictability: measures how hard it is to predict X_i from X_1, X_2, \dots, X_{i-1} (for $i \leftarrow [k]$).

Claim 19

Assume f is OWF, then $g(U_n, H_n)$ has next-block pseudoentropy $n + |h| + 1$.

Proof: Define $g'(h, x)$ by $g'(h, x)_i = \begin{cases} U, & i = n + |h| + d_f(x) + \log n \\ g(h, x)_i, & \text{otherwise.} \end{cases}$
 $g'(U_n, H_n)$ realizes the next-block pseudoentropy of $g(U_n, H_n)$.

Continue to Power-point presentation.